



Full Length Article

Cloning of *cbb* Gene Promoters from the Sulfur-Oxidizing Bacteria *Thiobacillus thioparus* and *Halothiobacillus neapolitanus*, and Assessment of their Promoter Efficacies

Xiao-Hua Fu^{1,2}, Ya-Nan Wang³, Yiu Fai Tsang⁴, Huan Li^{1,2}, Yi-Quan Le^{1,2} and Lei Wang^{1,2*}

¹State Key Laboratory of Pollution Control and Resource Reuse, College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China

²Shanghai Institute of Pollution Control and Ecological Security, China

³College of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao 266033, P.R. China

⁴Department of Science and Environmental Studies, The Education University of Hong Kong, Tai Po, New Territories, Hong Kong SAR, China

*For correspondence: celwang@tongji.edu.cn

Abstract

To determine the transcriptional potential of *cbb* genes involved in CO₂ fixation, the promoters of two sulfur-oxidizing bacteria, *Thiobacillus thioparus* DSM 505 and *Halothiobacillus neapolitanus* DSM 15147 were cloned. The unknown upstream sequence of the DSM 505 *cbbL* gene was successfully amplified through three rounds of thermal asymmetric interlaced PCR, yielding the fragments UP_{505-cbbL-Short} and UP_{505-cbbL-Long}. Sequence analysis confirmed that these two fragments contained the upstream sequence of the DSM 505 *cbbL* gene and that UP_{505-cbbL-Long} contained a promoter sequence. A series of fragments containing potential DSM 505 *cbbL* promoters (P_{505-cbbL-Y+1}, P_{505-cbbL-Y+1}, P_{505-cbbL-Y+1+1}, P_{505-cbbL-J+1} and P_{505-cbbL-J+1+1}) and fragments containing the known promoters of DSM 15147 *cbb* genes were inserted into the promoter-reporter plasmid pUC18-GFP, which contained the green fluorescent protein (GFP) gene. Sequencing results confirmed that all the promoter fragments were successfully cloned and transformed into *Escherichia coli* Top10. The promoter efficacies of the promoter fragments were then assessed by measuring the fluorescence intensities of the recombinant strains. For DSM 15147, the promoter efficacy of the *cbbL* promoter was much higher than that of the *cbbM* promoter, which corresponded with the transcription efficiencies of the *cbb* genes. For DSM 505, the promoter fragments P_{505-cbbL-Y+1+1} and P_{505-cbbL-J+1+1} were able to initiate GFP expression. The promoter efficacy of the DSM 505 *cbbL* promoter was higher than that of the DSM 15147 *cbbL* promoter, indicating that the *cbbL* gene transcription and CO₂ assimilation potentials of DSM 505 may be higher than those of DSM 15147. © 2019 Friends Science Publishers

Keywords: *cbb* Promoter; CO₂ assimilation potential; Promoting efficacy; Sulfur-oxidizing bacteria; Thermal asymmetric interlaced PCR

Introduction

Biological fixation of CO₂ holds great potential for mitigating global warming (López *et al.*, 2013). Chemoautotrophic bacteria have attracted significant attention due to their ability to fix CO₂ in the dark and in challenging environments such as barren desert soils and large-scale industrial waste gas treatment reactors (Saini *et al.*, 2011). Sulfur-oxidizing bacteria (SOB) are a major category of chemoautotrophic bacteria that are able to oxidize certain sulfur-containing compounds (*e.g.*, S²⁻, S⁰ and S₂O₃²⁻) for use as energy sources and to utilize CO₂ as their sole carbon source for growth (Baker *et al.*, 1998). Therefore, SOB are not only able to remove sulfur-containing pollutants from wastewater and industrial waste

gases (Chung *et al.*, 1998; Tóth *et al.*, 2015), but also have the capability to fix CO₂, even at high CO₂ concentrations (Baker *et al.*, 1998; Wang *et al.*, 2016b).

The Calvin-Benson-Bassham (CBB) pathway is considered to be the primary carbon assimilation pathway (Badger and Bek, 2008; Berg, 2011). Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO; encoded by the *cbb* gene) is a key enzyme in the CBB cycle (Tabita *et al.*, 2007, 2008). There are four forms of RuBisCO and most chemoautotrophic bacteria use form I and/or II, encoded by the *cbbL* and *cbbM* genes, respectively (Kusian and Bowien, 1997), to assimilate CO₂ via the CBB cycle (Berg, 2011). RuBisCO forms I and II contain key active sites in their large subunits and many SOB species use RuBisCO form I and/or II to fix CO₂ (Hernandez *et al.*, 1996; Baker *et al.*, 1998).

The type and transcription efficiency of the *cbb* gene is the major determining factor of carbon fixation efficiency (Tabita, 1988; Yoshizawa et al., 2004; Badger and Bek, 2008). However, *cbb* gene transcription is highly sensitive to environmental conditions, such as CO₂ concentration and O₂ content (Yoshizawa et al., 2004; Toyoda et al., 2005). Additionally, endogenous and exogenous organic matter can repress *cbb* gene transcription, inhibiting the CO₂ assimilation process (Friedrich, 1982; Kusian and Bowien, 1997; Hu et al., 2011; Wang et al., 2017, 2018). Different bacteria have different sensitivities to environmental conditions and organic matter inhibition (Kusian and Bowien, 1997; Hu et al., 2011; Wang et al., 2017, 2018). Therefore, the apparent gene transcription efficiency of *cbb* does not necessarily represent actual RuBisCO enzyme activity or CO₂ assimilation potential in chemoautotrophs.

Cloning of the gene encoding the RuBisCO enzyme from autotrophic bacteria and expression in *Escherichia coli* have been reported (Hernandez et al., 1996). However, expression levels of RuBisCO protein in heterotrophic *E. coli* are extremely low and the enzyme is generally inactive, which makes it difficult to determine the actual enzyme expression potential of RuBisCO (Kusian et al., 1995; Vichivanives et al., 2000). Therefore, analysis of the expression potential of the RuBisCO enzyme in chemoautotrophs has rarely been reported. However, understanding the expression potential of the RuBisCO enzyme is useful for screening bacteria with high CO₂ assimilation potentials and optimizing cultivation conditions to achieve high efficiency carbon fixation.

Promoter strength is the main factor influencing the transcription potential of structural genes (Dubbs and Tabita, 2003, 2004). Therefore, *cbb* gene promoter strength indirectly represents the transcription potential of *cbb* genes. The green fluorescent protein (GFP) gene has been widely used as a reporter for assessing promoter activity in prokaryotic organisms (Ikeno et al., 2003; Wei et al., 2012). However, *cbb* gene promoters in chemoautotrophs have rarely been studied and assessment of the expression potential of the RuBisCO enzyme in chemoautotrophs using promoter cloning technology has not been reported.

Our previous study showed that *Thiobacillus thioiparus* DSM 505 and *Halothiobacillus neapolitanus* DSM 15147 had similar high CO₂ fixation efficiencies but that the transcription efficiency of the DSM 505 *cbb* gene was much lower than that of the DSM 15147 *cbb* gene (Wang et al., 2016a). DSM 505 was also found more sensitive to inhibition by extracellular dissolved organic carbon than DSM 15147 (Wang et al., 2018). These results raised the question of which strain has a higher CO₂ assimilation potential under uninhibited conditions and should be optimized. To answer this question, the promoters of DSM 505 *cbbL* and DSM 15147 *cbb* genes were cloned and the promoter efficacies of the *cbb* promoters were measured to compare the *cbb* gene transcription potential of the two strains. The study will be helpful identify bacteria with high

CO₂ assimilation potentials and optimizing cultivation conditions to improve CO₂ fixation efficiency.

Materials and Methods

Bacterial Strains, Plasmids and Culture Conditions

Two aerobic SOB (*T. thioiparus* DSM 505 and *H. neapolitanus* DSM 15147) were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). *T. thioiparus* DSM 505 and *H. neapolitanus* DSM 15147 are obligate autotrophic bacteria, belonging to β -Proteobacteria and γ -Proteobacteria, respectively, that possess both *cbbL* and *cbbM* genes (Wang et al., 2016a). The strains were cultured using the method described by Wang et al. (2016a).

The plasmids used in this study are listed in Table 1. *E. coli* Top10 was used as the host strain for all recombinant plasmids. *E. coli* Top10 was grown at 37°C in Luria broth (LB) and strains containing plasmids were grown in LB supplemented with 60 μ g/mL ampicillin.

DNA Extraction

Genomic DNA was extracted from DSM 505 and DSM 15147 collected with a 0.22 μ m membrane. Total bacterial DNA was extracted from the filter membranes using a PowerSoil® DNA Isolation kit (Mo Bio Laboratories Inc., Carlsbad, C.A., U.S.A.) according to the manufacturer's protocol. The concentration and purity of DNA were measured using a NanoDrop ND-2000 spectrophotometer (Thermo Fisher Scientific, Waltham, M.A., U.S.A.).

Primers and PCR Procedures for Amplification of *cbb* Promoters

Because the full sequence of the DSM 505 *cbb* gene and its upstream sequence were unknown, a thermal asymmetric interlaced PCR (TAIL-PCR) was used to amplify the unknown upstream sequences (Liu and Chen, 2007; Qiu et al., 2008). The promoters of the DSM 505 *cbbL* gene were amplified using the high fidelity PCR components and procedures (Table 2 and 3). For DSM 15147, the upstream sequences of the *cbbL* (GenBank: AF038430.1) and *cbbM* (GenBank: AF046932.1) genes were acquired from the National Center for Biotechnology Information (NCBI). The promoters of the DSM 15147 *cbbL* and DSM 15147 *cbbM* genes were amplified using the conventional PCR components and procedures (Table 2 and 3). Primers used to amplify the *cbb* gene promoters and the GFP gene are shown in Table 4.

Construction of Recombinant Plasmids

The recombinant plasmid pUC18-GFP-P_{*cbb*} containing the DSM 505 and DSM 15147 *cbb* gene promoters were constructed (Fig. 1). Gene cloning and transformation were

Table 1: Main plasmids used in this study

Plasmids	Properties	Sources
pUC18	pBR322 ori, Amp ^R , Lac Z promoter, 2.7 kb	Lab stock
pMD18T	ColE1 ori, Amp ^R , Lac Z operater, 2.7 kb	TaKaRa
pET28b-GFP(+)	pET28b derivative, Kan ^R , containing the GFP gene, between BamH I and Hind III sites, 6.1 kb	Lab stock
pUC18-GFP	pUC18 derivative, containing the GFP gene, between BamH I and Hind III sites, 3.4 kb	This study
pUC18-GFP-P _{15147-cbbL}	pUC18-GFP derivative, containing the P _{15147-cbbL} gene, between Sac I and BamH I sites, 4.7 kb	This study
pUC18-GFP-P _{15147-cbbM}	pUC18-GFP derivative, containing the P _{15147-cbbM} gene, between Sac I and BamH I sites, 4.5 kb	This study
pUC18-GFP-P _{505-cbbL-Y+1}	pUC18-GFP derivative, containing the P _{505-cbbL-Y+1} gene, between Sac I and BamH I sites, 4.9 kb	This study
pUC18-GFP-P _{505-cbbL-Y+1}	pUC18-GFP derivative, containing the P _{505-cbbL-Y+1} gene, between Sac I and BamH I sites, 4.9 kb	This study
pUC18-GFP-P _{505-cbbL-Y+1+1}	pUC18-GFP derivative, containing the P _{505-cbbL-Y+1+1} gene, between Sac I and BamH I sites, 4.9 kb	This study
pUC18-GFP-P _{505-cbbL-J+1}	pUC18-GFP derivative, containing the P _{505-cbbL-J+1} gene, between Sac I and BamH I sites, 4.6 kb	This study
pUC18-GFP-P _{505-cbbL-J+1}	pUC18-GFP derivative, containing the P _{505-cbbL-J+1} gene, between Sac I and BamH I sites, 4.6 kb	This study
pUC18-GFP-P _{505-cbbL-J+1+1}	pUC18-GFP derivative, containing the P _{505-cbbL-J+1+1} gene, between Sac I and BamH I sites, 4.6 kb	This study

Amp: ampicillin; Kan: kanamycin

Table 2: PCR components for *cbb* promoters (25 μL)

PCR types	PCR components	Volume (μL)
Conventional-PCR (for DSM 15147)	10×PCR Buffer (Mg ²⁺ Plus)	2.50
	dNTP Mixture (2 mM each)	1.20
	Primer-F (10 μM)	0.60
	Prime -R (10 μM)	0.60
	rTaq DNA Polymerase	0.125
	ddH ₂ O	18.975
	DNA template	1.00
	TAIL-PCR (for DSM 505) (Qiu <i>et al.</i> , 2008)	10×PCR Buffer (Mg ²⁺ Plus)
dNTP Mixture (2 mM each)		1.20
12N Primer (10 μM)		4.00
Specific Primer (SP1/SP2/SP3) (10 μM)		1.00
rTaq DNA Polymerase		0.25
ddH ₂ O		15.05
High-fidelity PCR (for DSM 505)	DNA template	1.00
	5×PS Buffer	5.00
	dNTP Mixture(2 mM each)	2.00
	Primer-F (10 μM)	0.50
	Primer-R (10 μM)	0.50
	Primer HS DNA polymerase	0.25
	ddH ₂ O	15.75
	DNA template	1.00

Table 3: PCR procedure for *cbb* promoters

PCR procedure	Stage	Cycles	Standard thermal conditions
Conventional-PCR (for DSM 15147)	1	1	94°C (3 min)
	2	35	94°C (30 s), 55-56°C (30 s), 72°C (90s, adjusted according to the fragment length, 60s for 1 kb)
	3	1	72°C (7 min)
TAIL-PCR (for DSM 505) (Qiu <i>et al.</i> , 2008)	Primary	1	92°C (3 min), 95°C (1 min)
		2	94°C (30 s), 58°C (1 min), 72°C (2 min)
		3	94°C (30 s), 25°C (3 min), ramping to 72°C over 3min, 72°C (2 min)
		4	94°C (30 s), 29°C (1 min), 72°C (2 min)
		5	94°C (30 s), 58°C (1 min), 72°C (2 min); 94°C (30 s), 58°C (1 min), 72°C (2 min); 94°C (30 s), 29°C (1 min), 72°C (2 min)
	Secondary	1	72°C (5 min)
		1	94°C (30 s), 59°C (1 min), 72°C (2 min); 94°C (30 s), 59°C (1 min), 72°C (2 min); 94°C (30 s), 29°C (1 min), 72°C (2 min)
	Tertiary	2	72°C (5 min)
		1	94°C (30 s), 58°C (1 min), 72°C (2 min); 94°C (30 s), 58°C (1 min), 72°C (2 min); 94°C (30 s), 29°C (1 min), 72°C (2 min)
		2	72°C (5 min)
High fidelity-PCR (for DSM 505)	1	1	95°C (1 min)
	2	32	98°C (10 s), 56°C (10 s), 72°C (90s, adjusted according to the fragment length, 60s for 1 kb)
	3	1	72°C (7 min)

performed using PCR amplification, PCR product purification, restriction endonuclease digestion, agarose gel electrophoresis, recovery and purification of agarose gel

products, ligation, transformation, transformant PCR identification and sequencing. The reagents used for PCR amplification (TaKaRa, Dalian, China) are listed in Table 2.

Table 4: Primers used in this study

Target fragment	Primers	Sequence (5'-3')	Restriction sites	Target fragment length (bp)	
GFP gene	G1f	<u>CGCGGATCC</u> GATGGTGAGCAAGGGCGAG	BamH I	720	
	G720r	CCCA <u>AGCTTT</u> TACTTGTACAGCTCGTCCA	Hind III		
Insertion fragments in recombinant plasmids	M13F	CGCCAGGGTTTCCAGTCACGAC	/	/	
	M13R	AGCGGATAACAATTCACACAGGA	/	/	
P _{15147-cbbL}	hnLF	<u>TCCGAGCTC</u> CACACCGGGAAGTTTGGA	Sac I	1253	
	hnLR	<u>CGCGGATCC</u> TAGTCCATGTCGGTCAGCAA	BamH I		
P _{15147-cbbM}	hnMF	<u>TCCGAGCTC</u> TGATCGCTCGTFACTGTTC	Sac I	1072	
	hnMR	<u>CGCGGATCC</u> CGCATGTCTTCAGTGGCTTC	BamH I		
505-cbbL gene	K2F	ACCA <u>YCAAG</u> CCSAAGCTSGG	/	495	
	V2F	GCCTTCSAGCTTGCCSACCRC	/		
TAIL-PCR (DSM 505)	Primary	12N	NNNNNNNNNNNN	/	/
		V2F (SP1)	GCCTTCSAGCTTGCCSACCRC	/	/
	Secondary	12N	NNNNNNNNNNNN	/	/
		505L-R-II (SP2)	CTTCAGGGTTCGCTTCGTGC	/	/
	Tertiary	12N	NNNNNNNNNNNN	/	/
		505L-R-I (SP3)	<u>CGCGGATCC</u> GTTCCTCGTCCTTGGTG	BamH I	/
P _{505-cbbL-Y+1}	505L-F-YD	<u>TCCGAGCTC</u> ACTGTCCGGGCTGGTTGAGG	Sac I	1518	
	505L-R-I(-1)	<u>CGCGGATCC</u> TTCGTCGTCCTTGGTG	BamH I		
P _{505-cbbL-Y-1}	505L-F-YD	<u>TCCGAGCTC</u> ACTGTCCGGGCTGGTTGAGG	Sac I	1519	
	505L-R-I	<u>CGCGGATCC</u> GTTCCTCGTCCTTGGTG	BamH I		
P _{505-cbbL-Y+1+1}	505L-F-YD	<u>TCCGAGCTC</u> ACTGTCCGGGCTGGTTGAGG	Sac I	1520	
	505L-R-I(+1)	<u>CGCGGATCC</u> CGTTCTCGTCGTCCTTGGTG	BamH I		
P _{505-cbbL-J+1}	505L-F-J	<u>TCCGAGCTC</u> TGGCGAGCACCTGTTCTTGTG	Sac I	1165	
	505L-R-I(-1)	<u>CGCGGATCC</u> TTCGTCGTCCTTGGTG	BamH I		
P _{505-cbbL-J-1}	505L-F-J	<u>TCCGAGCTC</u> TGGCGAGCACCTGTTCTTGTG	Sac I	1166	
	505L-R-I	<u>CGCGGATCC</u> GTTCCTCGTCGTCCTTGGTG	BamH I		
P _{505-cbbL-J+1+1}	505L-F-J	<u>TCCGAGCTC</u> TGGCGAGCACCTGTTCTTGTG	Sac I	1167	
	505L-R-I(+1)	<u>CGCGGATCC</u> CGTTCTCGTCGTCCTTGGTG	BamH I		

Note: a) M=A/C; R = A/G; W=A/T; S= G/C; Y=C/T; K=G/T; V= A/G/C; H= A/C/T; D = A/G/T; B=G/C/T; N =A/G/C/T; b) The bases with underline are restriction sites and the italics are protect bases; c) The target fragment includes the promoter region and part of *cbb* gene, and the target fragment length includes the restriction sites

QuickCut *Bam*HI, QuickCut *Hind*III and QuickCut *Sac*I restriction enzymes (TaKaRa) were used for restriction endonuclease digestion. T4 DNA Ligase (TaKaRa) was used for ligation reactions. Purification of agarose gel and PCR products was performed using a gel extraction kit (Biodev-Tech, Beijing, China) according to the manufacturer's instructions. Nucleotide sequences were determined using a 3730 Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA). Plasmid DNA was isolated from *E. coli* host cells using a Plasmid Miniprep kit (Axygen, Union City, CA, USA).

Construction of Promoter-reporter Plasmid pUC18-GFP

The pET28b-GFP plasmid was maintained in our laboratory (State Key Laboratory of Pollution Control and Resource Reuse, Tongji University, Shanghai, China). The GFP gene fragment of pET28b-GFP and the pUC18 plasmid were digested with *Bam*HI and *Hind*III restriction enzymes, purified and ligated using T4 DNA Ligase. The ligation product was transformed into *E. coli* Top10. Transformants were randomly selected for PCR identification using M13F and M13R universal primers (Table 4) and positive transformants were sequenced to confirm that the inserted promoter sequences were correct. The recombinant plasmid pUC18-GFP was extracted using a Plasmid UP Midiprep Purification System B kit (Biodev-Tech) and used for follow-up experiments.

Cloning of the DSM 505 *cbb* Promoter

The unknown upstream fragment of DSM 505 *cbbL* was amplified using TAIL-PCR (Fig. 2). First, a fragment of the DSM 505 *cbbL* gene (S-DSM 505-*cbbL*) was amplified with primers K2F and V2F. The PCR product was gel purified and ligated with the pMD18-T plasmid (TaKaRa) and the resulting vector was transformed into *E. coli* Top10. The transformants were identified by PCR and sequenced. Second, three specific reverse primers were designed based on the S-DSM 505-*cbbL* sequence (Table 4). The unknown upstream sequence was amplified with the random primer 12N and three specific reverse primers using three rounds of TAIL-PCR (Table 2 and 3). The product of each round of PCR was used as the template for subsequent round of PCR. Upstream fragments were obtained after the third round of TAIL-PCR. After gel purification, the fragments were ligated with the pMD18-T plasmid and the resulting vectors were transformed into *E. coli* Top10. The inserted fragments were identified by PCR, sequenced and named UP_{505-cbbL}.

Because the exact location of the initiation codon of the *cbbL* gene was unknown, two specific reverse primers were designed to amplify products either one base longer [505L-R-I(+1)] or one base shorter [505L-R-I(-1)] than 505L-R-I. Additionally, to ensure that the upstream sequence contained the promoter region, two forward primers (505L-F-YD and 505L-F-J) were designed based on the UP_{505-cbbL} sequence (Table 4). Six putative promoter regions for DSM 505-*cbbL*

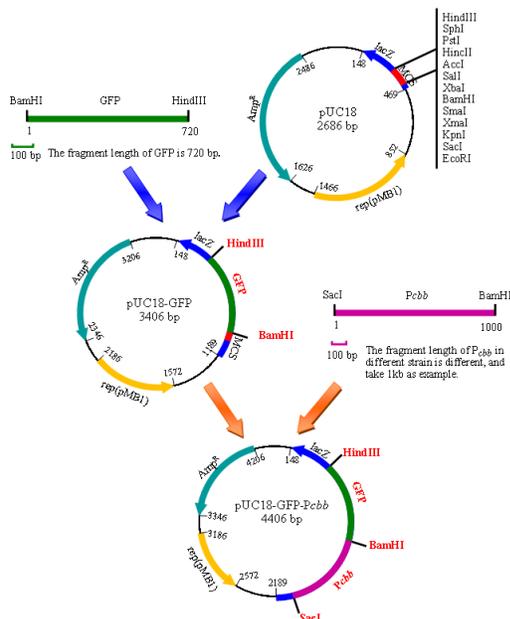


Fig. 1: Construction diagram of the recombinant plasmid pUC18-GFP-*P_{cbb}*. In which, the digested fragment of GFP gene was firstly inserted into the cloning vector pUC18 and formed a recombinant plasmid pUC18-GFP with *Bam*HI and *Hind*III restriction enzymes digestion. After that, the PCR amplification fragment of *cbb* gene promoter (*P_{cbb}*) was inserted into pUC18-GFP and formed a new recombinant plasmid pUC18-GFP-*P_{cbb}* with *Sac*I and *Bam*HI restriction enzymes digestion

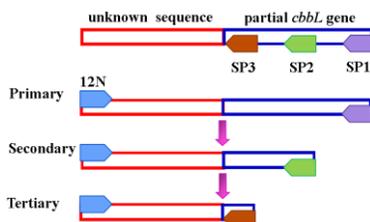


Fig. 2: Schematic diagram of amplifying the *cbbL* promoter in DSM 505 with unknown nucleotide sequence utilized thermal asymmetric interlaced PCR (TAIL-PCR) technology

were amplified by high-fidelity PCR and named *P_{505-cbbL-Y+I-1}*, *P_{505-cbbL-Y+I}*, *P_{505-cbbL-Y+I+I-1}*, *P_{505-cbbL-J+I-1}*, *P_{505-cbbL-J+I}* and *P_{505-cbbL-J+I+I-1}*. The PCR reaction products and the pUC18-GFP plasmid were digested with *Sac*I and *Bam*HI, ligated and transformed into *E. coli* Top10. The transformants were identified by PCR and sequenced. Among each set of three fragments, only the fragment containing the promoter with the correct sequence can initiate expression of GFP, which can be detected by measuring the fluorescence intensity of the recombinant *E. coli* strains.

Cloning the Known DSM 15147 *cbb* Promoter Sequence

The *cbbL* and *cbbM* promoters were first amplified using the primers hnLF/hnLR and hnMF/hnMR shown in Table 4. The

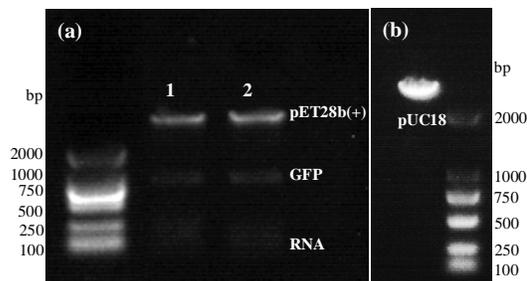


Fig. 3: Enzyme digestion results for plasmids pET28b-GFP and pUC18 with *Bam*HI and *Hind*III restriction enzymes digestion. Among them, the middle band of GFP fragment was obtained in (a) and the pUC18 band after digestion was obtained in (b). (Note: 1 and 2 are parallel samples)

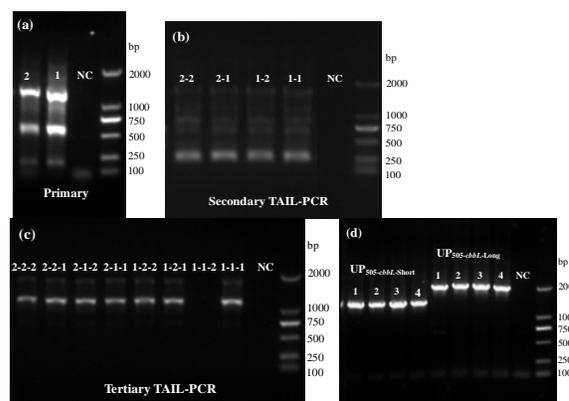


Fig. 4: Three rounds of TAIL-PCR amplification results for S-DSM 505-*cbbL* upstream sequence and the PCR identification results for transformants containing recombinant plasmid pMD18-UP505-*cbbL*-Long and pMD18-UP505-*cbbL*-Short, and their length were about 1.7 and 1.2 kb, respectively. (Note: NC represents the negative control, and 1, 2, 1-1, 1-2, 1-1-1, 1-1-2, 2-2-2 and 1,2,3,4 are parallel samples, respectively)

PCR reaction products *P_{15147-cbbL}* and *P_{15147-cbbM}* and the pUC18-GFP plasmid were digested with *Sac*I and *Bam*HI and ligated using T4 DNA Ligase. The ligation products pUC18-GFP-*P_{15147-cbbL}* and pUC18-GFP-*P_{15147-cbbM}* were transformed into *E. coli* Top10. Transformants were randomly selected for PCR identification using the universal primers M13F and M13R. Positive transformants were sequenced to confirm that the inserted promoter sequences were correct.

Determination of the Promoter Efficacy of the Cloned *cbb* Gene Promoter in *E. coli*

Recombinant strains were grown in LB liquid medium containing 60 µg/mL ampicillin. *E. coli* Top10 cultured in LB liquid medium without ampicillin was used as a negative control. The promoter efficacy was assayed by measuring the fluorescence intensity of cultured recombinant strains using a Synergy 4 multimode reader (BioTek Instruments, Winooski, VT, USA) at an excitation wavelength (*Ex*) of 470 nm and an emission wavelength of 510 nm. Spectral

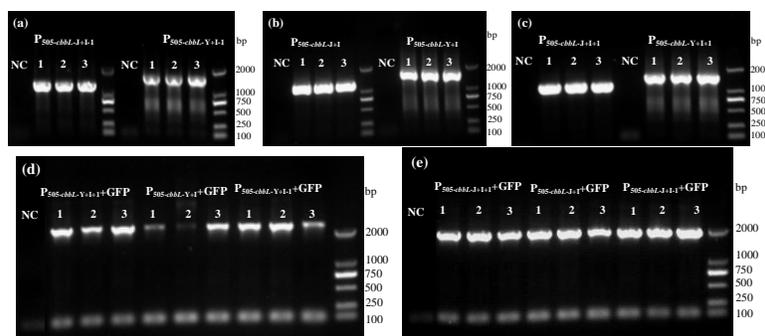


Fig. 5: PCR amplification results for the six fragments of DSM 505-*cbbL* gene upstream promoter (a, b and c); and PCR identification results for transformants containing their corresponding recombinant plasmids (d and e). The three long fragments of P_{505-cbbL-Y+I-1}, P_{505-cbbL-Y+I} and P_{505-cbbL-Y+I+1} were about 1.5 kb, and the three short fragments of P_{505-cbbL-J+I-1}, P_{505-cbbL-J+I}, and P_{505-cbbL-J+I+1} were about 1.2 kb. The long fragments of P_{505-cbbL-Y+I-1}+GFP, P_{505-cbbL-Y+I}+GFP and P_{505-cbbL-Y+I+1}+GFP were about 2.2 kb, and short fragments of P_{505-cbbL-J+I-1}+GFP, P_{505-cbbL-J+I}+GFP, and P_{505-cbbL-J+I+1}+GFP were about 1.9 kb. (Note: NC represents the negative control, and 1, 2 and 3 are parallel samples)

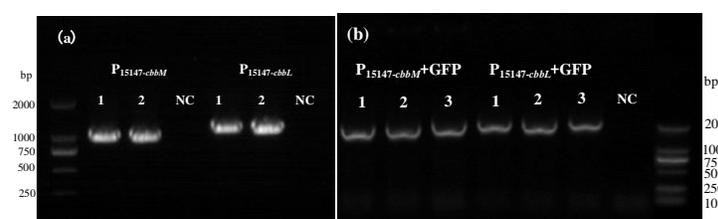


Fig. 6: PCR amplification results for *cbb* gene upstream promoter regions from *cbbL* and *cbbM* in DSM 15147 with fragments length of about 1.2 and 1.0 kb (a), and the PCR identification of transformants *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} and *E. coli* Top10/pUC18-GFP-P_{15147-cbbM} (b) with length of about 2.0 and 1.8 kb. (Note: NC represents the negative control, and 1, 2 and 3 are parallel samples)

scanning was also conducted at an *Ex* of 470 nm with an expected peak at 510 nm. To eliminate interference associated with LB medium background fluorescence, the medium was discarded and bacterial cells were washed three times with ddH₂O and suspended in ddH₂O.

Results

Construction of Promoter-report Plasmid pUC18-GFP

Double restriction enzyme-digested products were assessed by agarose gel electrophoresis. The results showed that the GFP gene fragment was excised from pET28b-GFP (Fig. 3a) and that the pUC18 plasmid was cut (Fig. 3b). PCR results showed that the GFP gene fragment was successfully inserted into the pUC18 plasmid. Results of transformant sequencing and sequence alignment showed that the sequence of the inserted GFP fragment was correct. Therefore, the recombinant bacteria *E. Coli* Top10/pUC18-GFP was obtained.

Cloning and Identification of DSM 505 *cbb* Promoter

A 500 bp fragment of the DSM 505 *cbbL* gene (S-DSM 505-*cbbL*) was amplified with primers K2F and V2F and cloned into the pMD18-T plasmid. PCR analysis of the transformants confirmed that the size of the inserted fragment was correct. Sequencing analysis showed that part

of the sequence at the 3'-end had 100% similarity with the NCBI Sequence Database sequence of DSM 505-*cbbL*-NCBI (gb DQ390449.1), indicating that the fragment is part of the DSM 505 *cbbL* gene. Because the cloned partial sequence aligned with the 5'-end of the NCBI sequence (gb DQ390449.1), we concluded that the sequence is located in the middle or upper regions of the *cbbL* gene. Three specific reverse primers for subsequent TAIL-PCR experiments were designed based on the obtained sequence (Fig. 2). The upstream sequence of S-DSM 505-*cbbL* was amplified through three rounds of TAIL-PCR (Fig. 2). Several fuzzy bands were observed after the first and second rounds of TAIL-PCR (Fig. 4a and b), whereas two legible bands, named UP_{505-cbbL-Short} and UP_{505-cbbL-Long}, were visible after the third round of TAIL-PCR (Fig. 4c). UP_{505-cbbL-Short} and UP_{505-cbbL-Long} were approximately 1.2 kb and 1.7 kb, respectively. The results of transformant PCR identification confirmed that the inserted fragments were the correct size (Fig. 4d).

Sequencing results were analysed using NCBI BLAST. Results showed that the 3' sequence of the 1148 bp UP_{505-cbbL-Short} was identical to the 5' sequence of S-DSM 505-*cbbL*, indicating that UP_{505-cbbL-Short} contains the upstream sequence of S-DSM 505-*cbbL*. Moreover, the full-length sequence of UP_{505-cbbL-Short} was identical to the 3' sequence of UP_{505-cbbL-Long}, indicating that the 1730 bp UP_{505-cbbL-Long}, which includes the full-length sequence of UP_{505-cbbL-Short}, contains a longer upstream sequence of S-DSM 505-*cbbL*. These results confirmed that both UP_{505-cbbL-Long} and UP_{505-cbbL-Short}

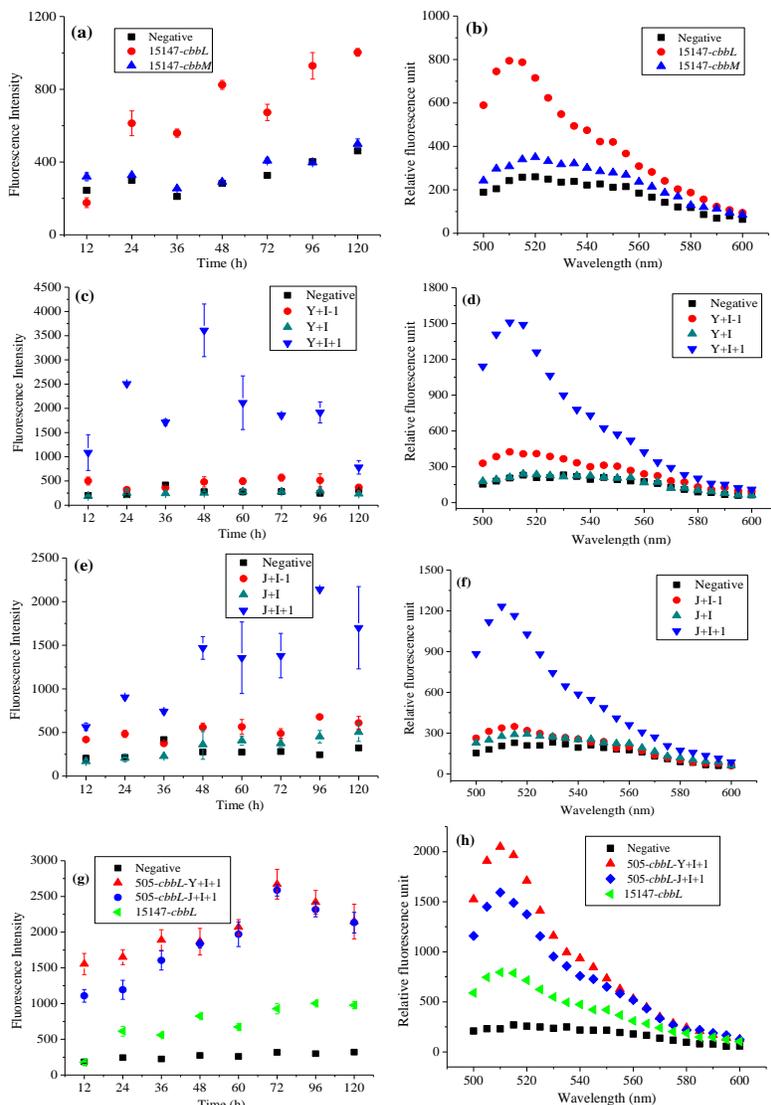


Fig. 7: Fluorescence intensity and emission spectra of recombinant bacteria in suspension at an excitation wavelength of 470 nm. (a): Fluorescence intensity of *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} and *E. coli* Top10/pUC18-GFP-P_{15147-cbbM}; (b): emission spectra of *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} and *E. coli* Top10/pUC18-GFP-P_{15147-cbbM} at 72 h during cultivation; (c)(d): Fluorescence intensity and emission spectra of the three recombinant bacteria strains containing long fragments of promoter (*E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I-1}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I} and *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1}); (e)(f): Fluorescence intensity and emission spectra of the three recombinant bacteria strains containing short fragments of promoter (*E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I-1}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I} and *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I+1}); (g)(h): the fluorescence intensity of the recombinant bacteria in suspension of *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I+1} and *E. coli* Top10/pUC18-GFP-P_{15147-cbbL}, and the excitation spectra at 96 h during cultivation

contained the correct upstream sequence of S-DSM 505-*cbbL*.

Based on sequence alignment of S-DSM 505-*cbbL* and DSM 505-*cbbL*-NCBI (gb DQ390449.1), and the fact that the length of the entire DSM 15147 *cbbL* gene was 1422 bp, we determined that S-DSM 505-*cbbL* may be located in either the 0–500 bp or 600–1000 bp region of the *cbbL* gene sequence. Because the upstream sequence of UP_{505-cbbL-Long} is 1730 bp, we determined that UP_{505-cbbL-Long} is located

approximately 1700 bp or 1100 bp upstream of the transcription start site of the DSM 505-*cbbL* gene. Therefore, UP_{505-cbbL-Long} must contain a promoter sequence. As shown in Fig. 5a, b, and c, six upstream promoter region fragments of DSM 505-*cbbL* (P_{505-cbbL-Y+I-1}, P_{505-cbbL-Y+I}, P_{505-cbbL-Y+I+1}, P_{505-cbbL-J+I-1}, P_{505-cbbL-J+I} and P_{505-cbbL-J+I+1}) were successfully PCR amplified.

Transformants containing each of the six fragments were PCR amplified with the universal primers M13F and

M13R (Fig. 5d and e). Since the inserted fragments include the GFP gene and the *cbb* promoter, the lengths of the target fragments were approximately 2.2 and 1.9 kb, as expected. Sequencing of transformants confirmed that the promoter sequences were correctly inserted into pUC18-GFP. Therefore, the recombinant strains *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I-1}, *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I} and *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I+1} were obtained.

Cloning and Identification of DSM 15147 *cbb* Promoter

The PCR reaction products P_{15147-cbbL} and P_{15147-cbbM} were visualized using agarose gel electrophoresis (Fig. 6a) and found to be approximately 1.2 kb and 1.0 kb, respectively, as expected. Transformant PCR identification results are shown in Fig. 6b. Since the inserted fragments include the GFP gene and the *cbb* promoter, the lengths of the target fragments were approximately 2.0 and 1.8 kb, as expected. Sequencing of positive transformants confirmed that the sequences of the inserted promoters and the open reading frame containing the GFP gene were correct. Thus, we expected the GFP gene to be expressed if the cloned promoter-containing fragment had promoter activity. Therefore, the recombinant bacteria *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} and *E. coli* Top10/pUC18-GFP-P_{15147-cbbM} were obtained.

Comparison of Promoter Efficacy of *cbb* Promoters in *E. coli*

The fluorescence intensities and emission spectra of the recombinant bacteria in suspension are shown in Fig. 7. For DSM 15147, an obvious emission peak at 510 nm was observed for *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} at an *Ex* of 470 nm. The fluorescence intensity of *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} was also significantly higher than that of *E. coli* Top10/pUC18-GFP-P_{15147-cbbM} (Fig. 7a and b).

For DSM 505, six recombinant strains containing different promoter fragments were constructed. Of the three recombinant bacteria containing fragments differing in sequence by one or two base pairs, only the one containing the promoter with the correct sequence could exhibit GFP fluorescence. Thus, the fluorescence intensities of the recombinant bacteria were measured to determine which promoter effectively initiated GFP expression. The fluorescence intensities and emission spectra of the recombinant bacteria in suspension at an *Ex* of 470 nm are shown in Fig. 7c-f. Obvious emission peaks at 510 nm were observed for *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1} and *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I+1} at an *Ex* of 470 nm. Higher fluorescence intensities were observed for *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1} and *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I+1}. And the fluorescence intensity of the recombinant bacteria *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1} was higher than that of *E. coli* Top10/pUC18-GFP-P₅₀₅₋

cbbL-J+I+1.

Fluorescence intensities of the recombinant bacteria *E. coli* Top10/pUC18-GFP-P_{505-cbbL-Y+I+1} and *E. coli* Top10/pUC18-GFP-P_{505-cbbL-J+I+1} in suspension were higher than those of *E. coli* Top10/pUC18-GFP-P_{15147-cbbL} (Fig. 7g and h). These results indicate that the promoter efficacy of the DSM 505 *cbbL* promoter was higher than that of the DSM 15147 *cbbL* promoter.

Discussion

In DSM 15147, the promoter strength of the *cbbL* gene was much higher than that of the *cbbM* gene. Similarly, the transcription efficiency of the *cbbL* gene was much higher than that of the *cbbM* gene (Wang *et al.*, 2016a), indicating that the *cbbL* gene might play a more significant role in CO₂ assimilation than the *cbbM* gene in DSM 15147.

For DSM 505, only the fragments amplified with the reverse primer 505-R-I (+1) contained the promoter with the correct sequence and were able to initiate GFP expression. These results were strengthened by the fact that both the long and short fragments amplified with the reverse primer 505-R-I (+1) were able to initiate GFP expression. The longer promoter fragments had higher promoter efficacy, this was likely due to the fact that the longer promoter fragments contain positive regulators; similar results had been reported for the *cbbL* promoter of *Rhodobacter capsulatus* (Paoli *et al.*, 1998).

The promoter efficacy of the DSM 505 *cbbL* promoter was higher than that of the DSM 15147 *cbbL* promoter, suggesting that the transcription potential of the DSM 505 *cbbL* gene should be higher than that of the DSM 15147 *cbbL* gene. However, the transcription efficiency of DSM 505 *cbbL* was lower than that of DSM 15147 *cbbL* (Wang *et al.*, 2016a). This indicated that the *cbbL* gene transcription and CO₂ assimilation potentials of DSM 505 might be higher than those of DSM 15147 but that the CO₂ assimilation potential of DSM 505 might not be fully realized.

The transcription activity of the *cbb* gene is known to be influenced by environmental factors (Yoshizawa *et al.*, 2004; Toyoda *et al.*, 2005), including CO₂ concentration (Wang *et al.*, 2016a; Wang *et al.*, 2016b) and exogenous organic matter (Friedrich, 1982; Kusian and Bowien, 1997; Wang *et al.*, 2017). Previously, it was found that DSM 505 was more sensitive to inhibition by extracellular dissolved organic carbon than DSM 15147 (Wang *et al.*, 2018). Therefore, optimization or elimination of such adverse factors could improve the transcription efficiency of the DSM 505 *cbbL* gene, resulting in higher CO₂ fixation efficiencies.

Conclusion

Regarding DSM 15147, the promoter efficacy of the *cbbL* promoter was much higher than that of the *cbbM* promoter, which corresponded with the *cbb* gene transcription

efficiency. However, in contrast to the actual expression efficiency, the promoter efficacy of the DSM 505 *cbbL* promoter was higher than that of DSM 15147. Our results suggest that the *cbbL* gene transcription and CO₂ assimilation potentials of DSM 505 are higher than those of DSM 15147 but that the gene transcription potential of DSM 505 *cbbL* is not fully realized.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 21577101), China National Key Research and Development Project (No. 2017YFC0506004), the Shandong Province Natural Science Foundation (No. ZR2018BEE037), and Major Project of the Shanghai Scientific and Technological Committee (16DZ1205903).

References

- Badger, M.R. and E.J. Bek, 2008. Multiple rubisco forms in proteobacteria: their functional significance in relation to CO₂ acquisition by the CBB cycle. *J. Exp. Bot.*, 59: 1525–1541
- Baker, S.H., S. Jin, H.C. Aldrich, G.T. Howard and J.M. Shively, 1998. Insertion mutation of the form I *cbbL* gene encoding ribulose biphosphate carboxylase/oxygenase (RuBisCO) in *Thiobacillus neapolitanus* results in expression of form II RuBisCO, loss of Carboxysomes, and an increased CO₂ requirement for growth. *J. Bacteriol.*, 180: 4133–4139
- Berg, I.A., 2011. Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl. Environ. Microbiol.*, 77: 1925–1936
- Chung, Y.C., C. Huang, J.R. Pan and C.P. Tseng, 1998. Comparison of autotrophic and mixotrophic biofilters for H₂S removal. *J. Environ. Eng.*, 124: 362–367
- Dubbs, J.M. and F.R. Tabita, 2004. Regulators of nonsulfur purple phototrophic bacteria and the interactive control of CO₂ assimilation, nitrogen fixation, hydrogen metabolism and energy generation. *FEMS Microbiol. Rev.*, 28: 353–376
- Dubbs, J.M. and F.R. Tabita, 2003. Interactions of the *cbbII* promoter-operator region with *cbbR* and *RegA* (PrrA) regulators indicate distinct mechanisms to control expression of the two *cbb* operons of *Rhodobacter sphaeroides*. *J. Biol. Chem.*, 278: 16443–16450
- Friedrich, C.G., 1982. Depression of hydrogenase during limitation of electron donors and derepression of ribulose biphosphate carboxylase during carbon limitation of *Alcaligenes eutrophus*. *J. Bacteriol.*, 149: 203–210
- Hernandez, J.M., S.H. Baker, S.C. Lorbach, J.M. Shively and F.R. Tabita, 1996. Deduced amino acid sequence, functional expression, and unique enzymatic properties of the form I and form II ribulose biphosphate carboxylase/oxygenase from the chemoautotrophic bacterium *Thiobacillus denitrificans*. *J. Bacteriol.*, 178: 347–356
- Hu, J.J., L. Wang, S.P. Zhang, Y.Q. Wang and X.F. Xi, 2011. Inhibitory effect of organic carbon on CO₂ fixing by non-photosynthetic microbial community isolated from the ocean. *Bioresour. Technol.*, 102: 7147–7153
- Ikeno, S., C. Ogino, T. Ito and N. Shimizu, 2003. Detection of benzene derivatives by recombinant *E. coli* with *Ps* promoter and GFP as a reporter protein. *Biochem. Eng. J.*, 15: 193–197
- Kusian, B. and B. Bowien, 1997. Organization and regulation of *cbb* CO₂ assimilation genes in autotrophic bacteria. *FEMS Microbiol. Rev.*, 21: 135–155
- Kusian, B., R. Bednarski, M. Husemann and B. Bowien, 1995. Characterization of the duplicate ribulose-1,5-bisphosphate carboxylase genes and *cbb* promoters of *Alcaligenes eutrophus*. *J. Bacteriol.*, 177: 4442–4450
- Liu, Y.G. and Y. Chen, 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques*, 43: 649–656
- López, J.C., G. Quijano, T.S. Souza, J.M. Estrada, R. Lebrero and R. Muñoz, 2013. Biotechnologies for greenhouse gases (CH₄, N₂O and CO₂) abatement: state of the art and challenges. *Appl. Environ. Microbiol.*, 97: 2277–2303
- Paoli, G.C., P. Vichivanives and F.R. Tabita, 1998. Physiological control and regulation of the *Rhodobacter capsulatus cbb* operons. *J. Bacteriol.*, 180: 4258–4269
- Qiu, Y., J. Tian, R. Ge, B. Zhao, Y. Shen and Z. Huang, 2008. A modified TAIL-PCR and its application in isolating gene promoter of wheat. *Chin. J. Biotechnol.*, 24: 695–699
- Saini, R., R. Kapoor, R. Kumar, T.O. Siddiqi and A. Kumar, 2011. CO₂ utilizing microbes - a comprehensive review. *Biotechnol. Adv.*, 29: 949–960
- Tóth, G., N. Nemestóthy, K. Bélafi-Bakó, D. Vozik and P. Bakonyi, 2015. Degradation of hydrogen sulfide by immobilized *Thiobacillus thioparus* in continuous biotrickling reactor fed with synthetic gas mixture. *Intl. Biodeter. Biodegrad.*, 105: 185–191
- Tabita, F.R., 1988. Molecular and cellular regulation of autotrophic carbon dioxide fixation in microorganisms. *Microbiol. Rev.*, 52: 155–189
- Tabita, F.R., S. Satagopan, T.E. Hanson, N.E. Kreele and S.S. Scott, 2008. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.*, 59: 1515–1524
- Tabita, F.R., T.E. Hanson, H. Li, S. Satagopan, J. Singh and S. Chan, 2007. Function, structure, and evolution of the RubisCO-like proteins and their Rubis CO homologs. *Microbiol. Mol. Biol. Rev.*, 71: 576–599
- Toyoda, K., Y. Yoshizawa, H. Arai, M. Ishii and Y. Igarashi, 2005. The role of two *cbbR*s in the transcriptional regulation of three ribulose-1,5-bisphosphate carboxylase/oxygenase genes in *Hydrogenovibrio marinus* strain MH-110. *Microbiology*, 151: 3615–3625
- Vichivanives, P., T.H. Bird, C.E. Bauer and F.R. Tabita, 2000. Multiple regulators and their interactions *in vivo* and *in vitro* with the *cbb* regulons of *Rhodobacter capsulatus*. *J. Mol. Biol.*, 300: 1079–1099
- Wang, Y.N., Y.F. Tsang, L. Wang, X. Fu, J. Hu, H. Li and Y. Le, 2018. Inhibitory effect of self-generated extracellular dissolved organic carbon on carbon dioxide fixation in sulfur-oxidizing bacteria during a chemoautotrophic cultivation process and its elimination. *Bioresour. Technol.*, 252: 44–51
- Wang, Y.N., Y.F. Tsang, L. Wang, X. Fu, H. Li, J. Hu and Y. Le, 2017. Influence of reduced sulfur on carbon fixation efficiency of *Halothiobacillus neapolitanus* and its mechanism. *Chem. Eng. J.*, 326: 249–256
- Wang, Y.N., L. Wang, Y.F. Tsang, X. Fu, J. Hu, H. Li and Y. Le, 2016a. Response of *cbb* gene transcription levels of four typical sulfur-oxidizing bacteria to the CO₂ concentration and its effect on their carbon fixation efficiency during sulfur oxidation. *Enzyme Microb. Technol.*, 92: 31–40
- Wang, Y.N., L. Wang, Y.F. Tsang, X. Fu, J. Hu, H. Li and Y. Le, 2016b. The variability in carbon fixation characteristics of several typical chemoautotrophic bacteria at low and high concentrations of CO₂ and its mechanism. *Intl. Biodeter. Biodegrad.*, 113: 105–112
- Wei, M., Z. Kong, L. Zhong, L. Qiu, Y. Li, L. Zhao, X. Li and W. Zhong, 2012. Construction of a native promoter-containing transposon vector for the stable monitoring of the denitrifying bacterium *Pseudomonas stutzeri* LYS-86 by chromosomal-integrated *gfp*. *Plasmid*, 68: 61–68
- Yoshizawa, Y., K. Toyoda, H. Arai, M. Ishii and Y. Igarashi, 2004. CO₂-responsive expression and gene organization of three ribulose-1,5-bisphosphate carboxylase/oxygenase enzymes and carboxysomes in *Hydrogenovibrio marinus* strain MH-110. *J. Bacteriol.*, 186: 5685–5691

[Received 06 May 2019; Accepted 23 May 2019; published (online) 10 Nov 2019]